Colorado Pediatric Vaccine Analysis

Brian Erly, MD MPH; Shannon O'Brien, MD MPH

2022-08-16

Introduction

The purpose of this analysis is to understand the factors which influence the uptake of the COVID vaccine among children and adolescents in Colorado.

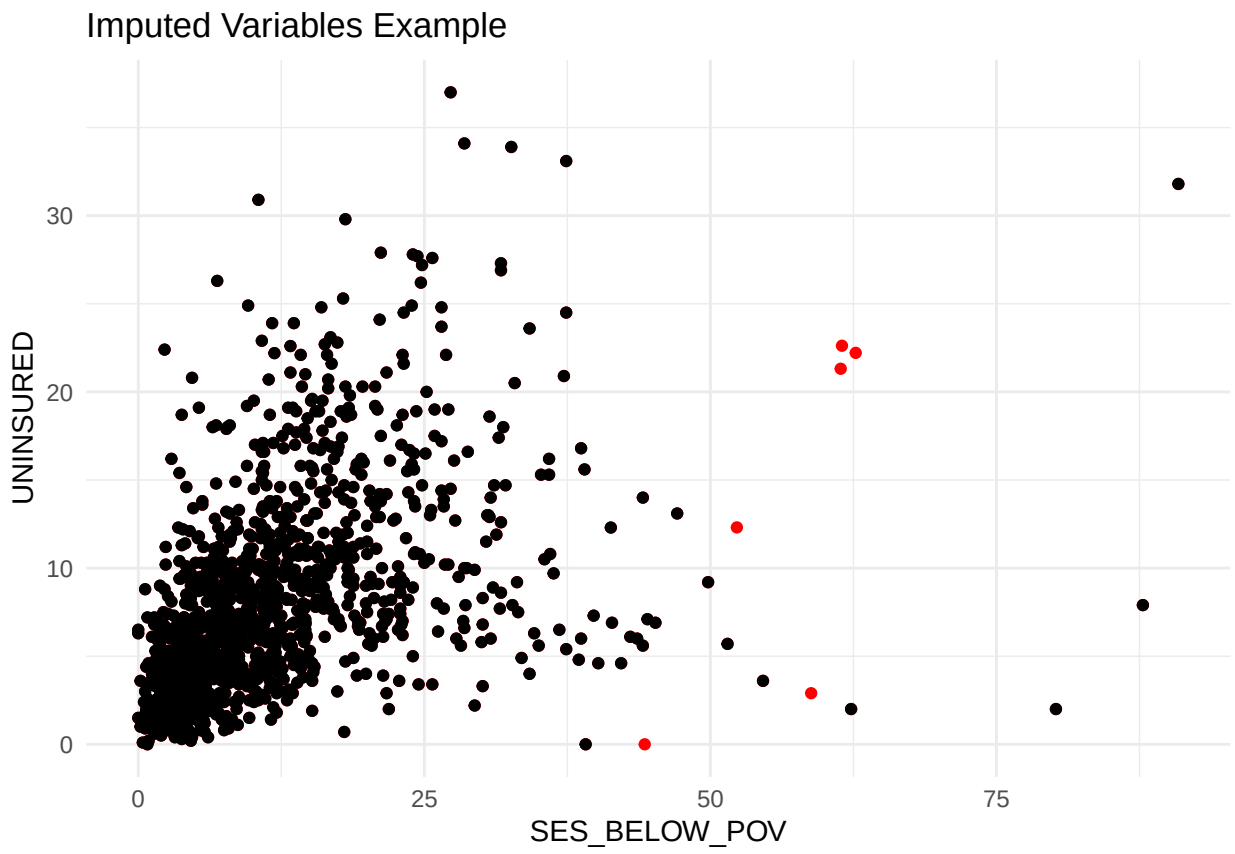
The factors which influence vaccine uptake can be conceptually divided into two categories. The first category are factors which are essentially non-modifiable (at least with respect to public health interventions). These include:
* Where a person lives and what their family income is
* Underlying health status
* What a person's racial/ethnic background is
* Whether a person has health insurance or lives near a hospital
* Where a person gets their news and information

Other modifiable factors might include
* Messaging received around vaccines
* Availability of transportation to the vaccine
* Workplace vaccine mandates or requirements
* Perceived risk of COVID-19

Identifying the underlying non-modifiable risk factors allows public health workers to first understand why vaccine uptake might be higher or lower in certain areas of the state. Next, we can identify over- or under-performing counties or regions to better understand the effects of modifiable risk factors. Finally, this information may help target interventions more effectively.

To approach this, we will use a combination of linear modeling and published frameworks for vaccine uptake to identify

```
## Warning: Removed 6 rows containing missing values (geom_point).
```

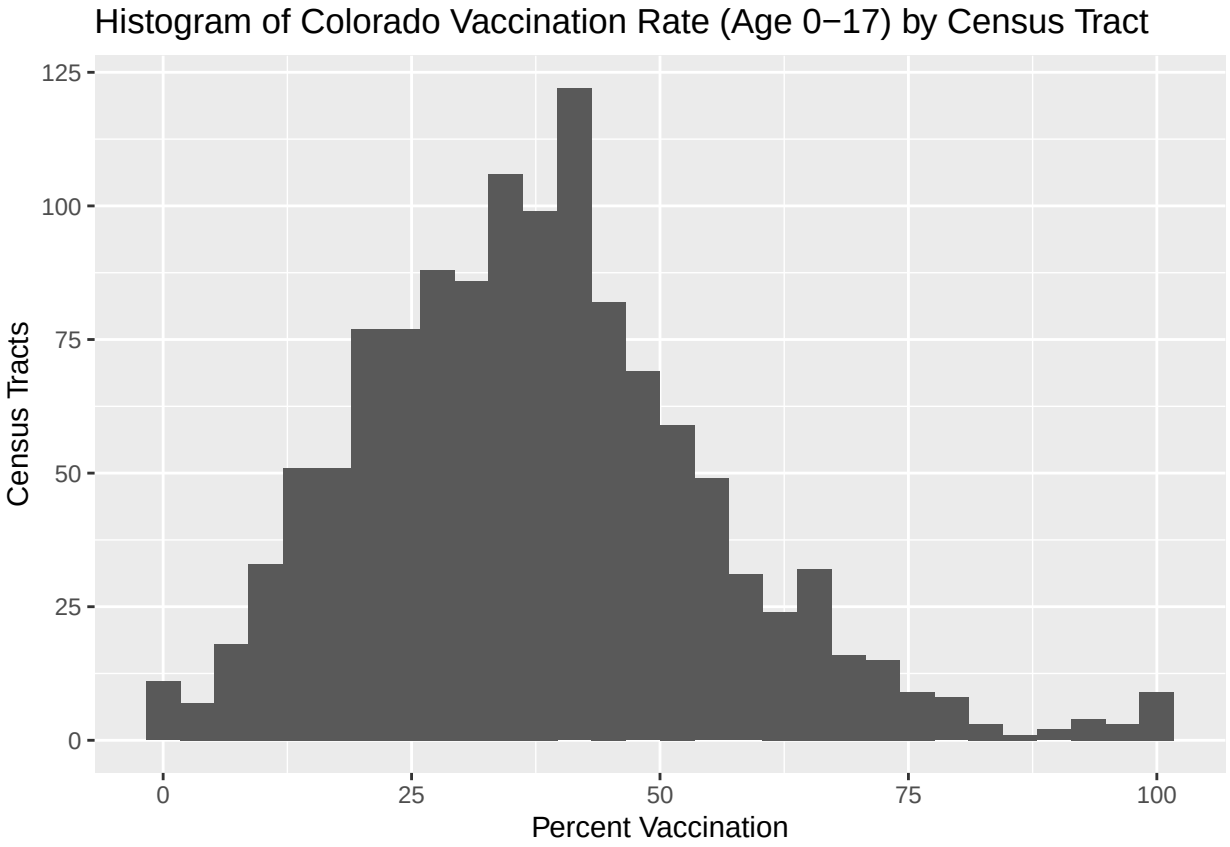



```
## Data for 1242 census tracts loaded.
```

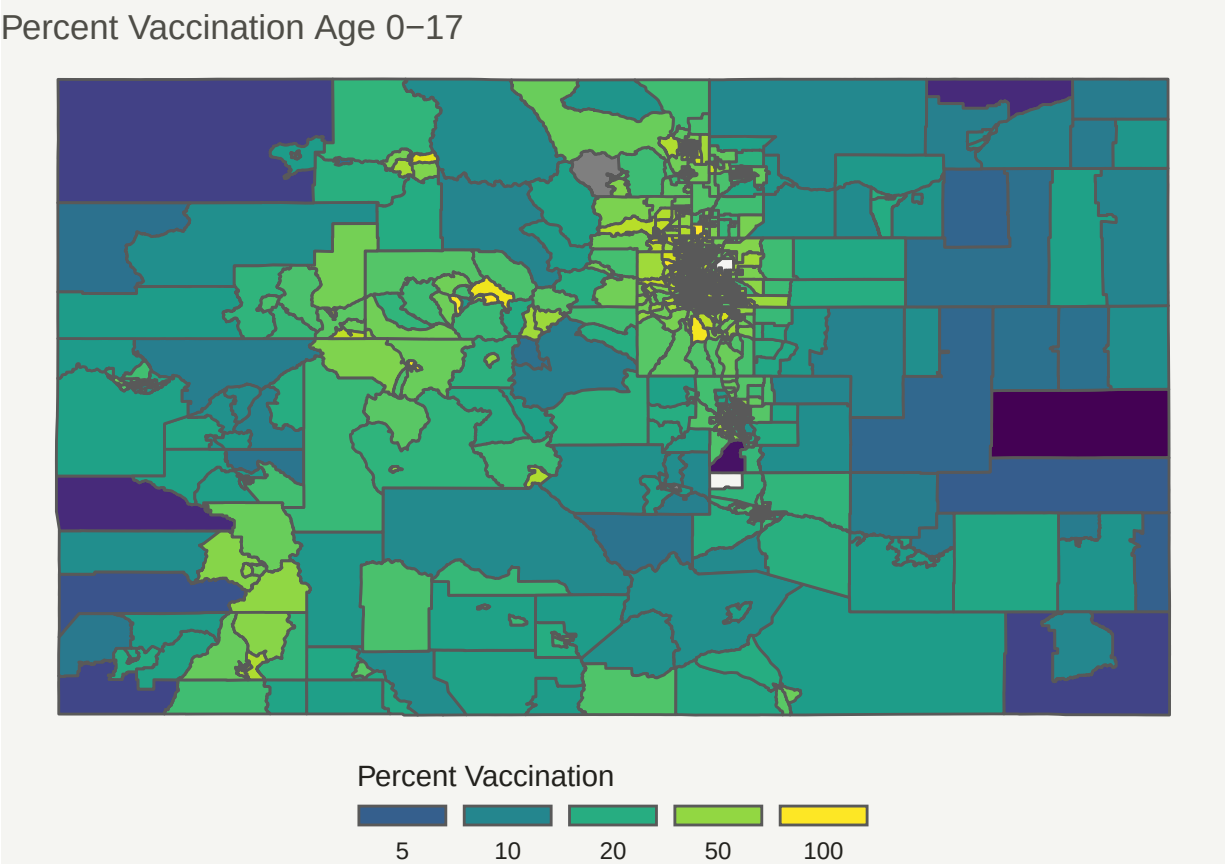
Basic Graphs

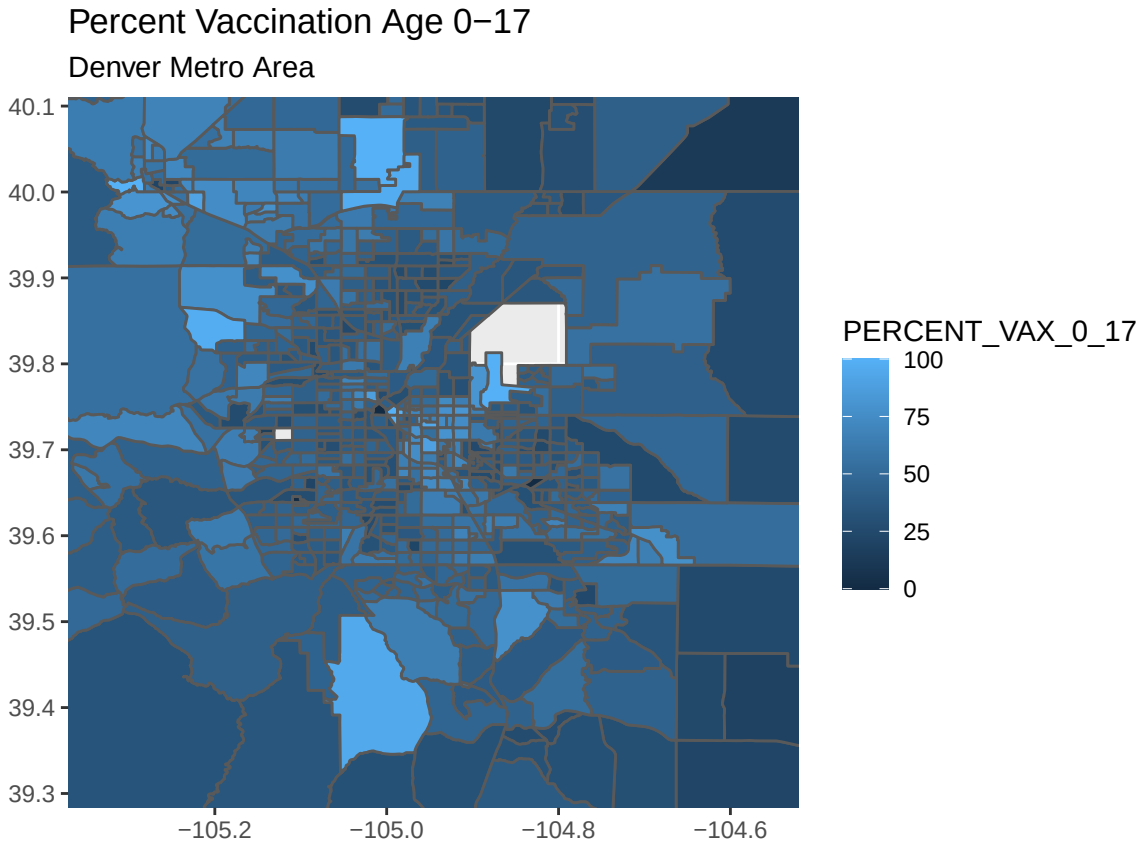
Data has been loaded. This includes the following data elements:

- COVID vaccination rates by age group and census tract
- COVID vaccination providers by census tract
- 2020 presidential election results
- Health Service Area hospital capacity
- Non-COVID vaccine rates by county for MMR, HPV, and 2019 Influenza
- Social Vulnerability Index data




```
## Warning: Transformation introduced infinite values in discrete y-axis
```


Percent Vaccination Age 0−17
Percent Vaccination
5
10
20
50
100



Linear Model

Using the available data, we can create a predicted vaccination rate for each census tract in Colorado using a linear model. This model creates a predicted vaccination rate based on non-modifiable factors for each census tract. By comparing the true vaccination rate with the predicted vaccination rate, we can identify under- and over-performing regions of the state.

The variables in the full model include the following:

- SES_BELOW_POV: Percent of individuals with household income below poverty line (from SVI)
- SES_UNEMPLOYED: Percent of individuals age 16+ who areunemployed (from SVI)
- SES_INCOME: Per capita income (from SVI)
- SES_NO_HS: Percent of individuals 25+ with no high school diploma (from SVI)
- HHCOMP_AGE65: Percent of individuals aged 65 and older (from SVI)
- HHCOMP_AGE17: Percent of individuals aged 17 and younger (from SVI)
- HHCOMP_DISABILITY: Percent of individuals with a disability (from SVI)
- HHCOMP_SING_PARENT: Percent of households with children under 18 with only one parent present (from SVI)

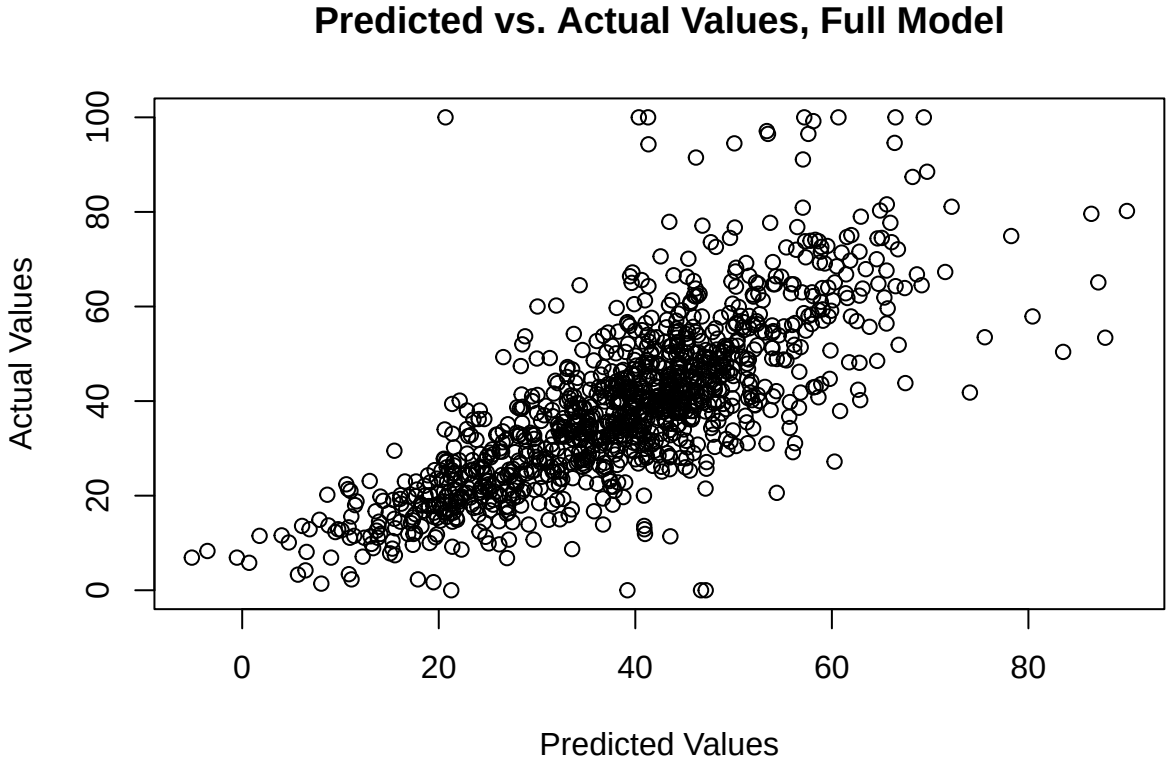
- MINORITY_MINORITY: Percent of individuals who are not nonhispanic White (from SVI)
- MINORITY_NONENGLISH: Percent of individuals age 5+ who speak English "less than well" (from SVI)
- HOUSING_MULTIUNIT: Percent of households in a structure with 10 or more households (from SVI)
- HOUSING_MOBILE: Percent of households that are mobile homes (from SVI)
- HOUSING_CROWDED: Percent of households with more people than rooms (from SVI)
- HOUSING_NO_VEHICLE: Percent of households with no vehicle (from SVI)
- HOUSING_GROUP_QUARTERS: Percent of persons in institutionalized group quarters (from SVI)
- UNINSURED: Percent of individuals who are uninsured (from SVI)
- MMR+: Percent of individuals aged 2-18 who are up-to-date on MMR vaccination (from CDPHE)
- HPV+: Percent of individuals aged 3-17 who received an HPV shot (from CDPHE)
- PedsFlu: Percent of individuals 17 and under who received an influenza vaccine in the 2019-2020 flu season (from CDPHE)
- AllFlu: Percent of individuals who received an influenza vaccine in the 2019-2020 flu season (from CDPHE)
- HSA_beds: Number of bed in the county Health Service Area (from SEER)
- pctBiden: Percent of individuals with household income below poverty line (from New York Times)
- providers_county: Number of COVID vaccine providers in the county per 1000 (from CDPHE)

A reduced model was also developed using backwards automated model selection.

```
## Warning: Removed 123 rows containing missing values.
```

```
##
## Call:
## lm(formula = PERCENT_VAX_0_17 ~ SES_BELOW_POV + SES_UNEMPLOYED +
##     SES_INCOME + SES_NO_HS + HHCOMP_AGE65 + HHCOMP_AGE17 + HHCOMP_DISABILITY +
##     HHCOMP_SING_PARENT + MINORITY_MINORITY + MINORITY_NONENGLISH +
##     HOUSING_MULTIUNIT + HOUSING_MOBILE + HOUSING_CROWDED + HOUSING_NO_VEHICLE +
##     HOUSING_GROUP_QUARTERS + UNINSURED + MMR + HPV + PedsFlu +
##     AllFlu + HSA_beds + pctBiden + providers_county, data = model_data,
##     weights = Under_18)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -888.06 -175.76  -21.66  163.74 2552.61
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           38.89144    0.32558 119.454  < 2e-16 ***
## SES_BELOW_POV          0.32506    0.60564   0.537  0.59157
## SES_UNEMPLOYED        -0.03974    0.41338  -0.096  0.92342
## SES_INCOME             7.34998    0.53628  13.706  < 2e-16 ***
```

```
## SES_NO_HS              -0.38095    0.71772  -0.531  0.59568
## HHCOMP_AGE65           -1.13297    0.56536  -2.004  0.04532 *
## HHCOMP_AGE17           -1.66921    0.57034  -2.927  0.00350 **
## HHCOMP_DISABILITY      -1.63379    0.60657  -2.693  0.00718 **
## HHCOMP_SING_PARENT     -1.18333    0.44921  -2.634  0.00855 **
## MINORITY_MINORITY       2.08231    0.70149   2.968  0.00306 **
## MINORITY_NONENGLISH     1.63934    0.65013   2.522  0.01182 *
## HOUSING_MULTIUNIT      -2.44745    0.51410  -4.761 2.19e-06 ***
## HOUSING_MOBILE          1.07743    0.34227   3.148  0.00169 **
## HOUSING_CROWDED        -1.22966    0.47963  -2.564  0.01049 *
## HOUSING_NO_VEHICLE      0.84558    0.53591   1.578  0.11489
## HOUSING_GROUP_QUARTERS  0.82634    0.58711   1.407  0.15957
## UNINSURED              -1.11546    0.51885  -2.150  0.03178 *
## MMR                    -0.90670    0.99527  -0.911  0.36249
## HPV                     2.55226    1.64738   1.549  0.12160
## PedsFlu                 0.56497    1.18456   0.477  0.63350
## AllFlu                  1.00883    0.78129   1.291  0.19690
## HSA_beds               -0.48026    0.60201  -0.798  0.42518
## pctBiden                5.94831    0.67557   8.805  < 2e-16 ***
## providers_county       -1.08198    0.59482  -1.819  0.06918 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 308.5 on 1092 degrees of freedom
## Multiple R-squared:  0.6371, Adjusted R-squared:  0.6295
## F-statistic: 83.36 on 23 and 1092 DF,  p-value: < 2.2e-16
```



```
## [1] "Now we will create a reduced model using backwards stepwise model selection."

## Start:  AIC=12853.94
## PERCENT_VAX_0_17 ~ SES_BELOW_POV + SES_UNEMPLOYED + SES_INCOME +
##     SES_NO_HS + HHCOMP_AGE65 + HHCOMP_AGE17 + HHCOMP_DISABILITY +
##     HHCOMP_SING_PARENT + MINORITY_MINORITY + MINORITY_NONENGLISH +
##     HOUSING_MULTIUNIT + HOUSING_MOBILE + HOUSING_CROWDED + HOUSING_NO_VEHICLE +
##     HOUSING_GROUP_QUARTERS + UNINSURED + MMR + HPV + PedsFlu +
##     AllFlu + HSA_beds + pctBiden + providers_county
##
##                           Df Sum of Sq       RSS   AIC
## - SES_UNEMPLOYED           1       879 103899001 12846
## - PedsFlu                  1     21643 103919765 12846
## - SES_NO_HS                1     26805 103924927 12846
## - SES_BELOW_POV            1     27408 103925530 12846
## - HSA_beds                 1     60554 103958675 12847
## - MMR                      1     78963 103977085 12847
## - AllFlu                   1    158631 104056753 12848
## <none>                                 103898122 12848
## - HOUSING_GROUP_QUARTERS   1    188479 104086601 12848
## - HPV                      1    228374 104126496 12848
## - HOUSING_NO_VEHICLE       1    236868 104134990 12848
## - providers_county         1    314813 104212935 12849
## - HHCOMP_AGE65             1    382099 104280221 12850
## - UNINSURED                1    439758 104337880 12851
## - MINORITY_NONENGLISH      1    604958 104503080 12852
## - HOUSING_CROWDED          1    625389 104523511 12853
## - HHCOMP_SING_PARENT       1    660220 104558342 12853
## - HHCOMP_DISABILITY        1    690251 104588373 12853
## - HHCOMP_AGE17             1    814961 104713082 12855
## - MINORITY_MINORITY        1    838364 104736486 12855
## - HOUSING_MOBILE           1    942812 104840934 12856
## - HOUSING_MULTIUNIT        1   2156389 106054511 12869
## - pctBiden                 1   7376253 111274374 12923
## - SES_INCOME               1  17872297 121770418 13024
##
## Step:  AIC=12851.95
## PERCENT_VAX_0_17 ~ SES_BELOW_POV + SES_INCOME + SES_NO_HS + HHCOMP_AGE65 +
##     HHCOMP_AGE17 + HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + MMR + HPV + PedsFlu + AllFlu + HSA_beds + pctBiden +
##     providers_county
##
##                           Df Sum of Sq       RSS   AIC
## - PedsFlu                  1     22133 103921134 12844
## - SES_NO_HS                1     26230 103925231 12844
## - SES_BELOW_POV            1     27011 103926012 12844
## - HSA_beds                 1     59809 103958810 12845
## - MMR                      1     78323 103977324 12845
## - AllFlu                   1    158215 104057216 12846
## <none>                                 103899001 12846
## - HOUSING_GROUP_QUARTERS   1    187683 104086684 12846
## - HPV                      1    227612 104126613 12846
```

```
## - HOUSING_NO_VEHICLE      1    236528 104135529 12846
## - providers_county        1    314258 104213259 12847
## - HHCOMP_AGE65            1    382616 104281618 12848
## - UNINSURED               1    439144 104338145 12849
## - MINORITY_NONENGLISH     1    604491 104503492 12850
## - HOUSING_CROWDED         1    625071 104524072 12851
## - HHCOMP_SING_PARENT      1    659818 104558820 12851
## - HHCOMP_DISABILITY       1    696487 104595488 12851
## - HHCOMP_AGE17            1    823804 104722805 12853
## - MINORITY_MINORITY       1    837739 104736740 12853
## - HOUSING_MOBILE          1    942302 104841303 12854
## - HOUSING_MULTIUNIT       1   2162094 106061095 12867
## - pctBiden                1   7400874 111299875 12921
## - SES_INCOME              1  17920641 121819642 13022
##
## Step:  AIC=12850.19
## PERCENT_VAX_0_17 ~ SES_BELOW_POV + SES_INCOME + SES_NO_HS + HHCOMP_AGE65 +
##     HHCOMP_AGE17 + HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + MMR + HPV + AllFlu + HSA_beds + pctBiden + providers_county
##
##                          Df Sum of Sq       RSS   AIC
## - SES_BELOW_POV           1     26705 103947839 12842
## - SES_NO_HS               1     27412 103948546 12842
## - HSA_beds                1     54550 103975684 12843
## - MMR                     1     77912 103999046 12843
## - HOUSING_GROUP_QUARTERS  1    183846 104104980 12844
## <none>                                103921134 12844
## - HOUSING_NO_VEHICLE      1    245925 104167059 12845
## - providers_county        1    311145 104232279 12846
## - HPV                     1    311193 104232327 12846
## - HHCOMP_AGE65            1    402756 104323890 12846
## - UNINSURED               1    435445 104356579 12847
## - AllFlu                  1    524766 104445900 12848
## - MINORITY_NONENGLISH     1    614461 104535595 12849
## - HOUSING_CROWDED         1    624047 104545181 12849
## - HHCOMP_SING_PARENT      1    656047 104577181 12849
## - HHCOMP_DISABILITY       1    741988 104663122 12850
## - MINORITY_MINORITY       1    816173 104737307 12851
## - HHCOMP_AGE17            1    831189 104752323 12851
## - HOUSING_MOBILE          1    928643 104849777 12852
## - HOUSING_MULTIUNIT       1   2160016 106081150 12865
## - pctBiden                1   7847318 111768452 12924
## - SES_INCOME              1  18112167 122033300 13022
##
## Step:  AIC=12848.47
## PERCENT_VAX_0_17 ~ SES_INCOME + SES_NO_HS + HHCOMP_AGE65 + HHCOMP_AGE17 +
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + MMR + HPV + AllFlu + HSA_beds + pctBiden + providers_county
##
##                          Df Sum of Sq       RSS   AIC
```

```
## - SES_NO_HS               1     18116 103965955 12841
## - HSA_beds                1     65333 104013172 12841
## - MMR                     1     79560 104027399 12841
## <none>                                103947839 12842
## - HOUSING_GROUP_QUARTERS  1    195000 104142839 12843
## - providers_county        1    298274 104246113 12844
## - HPV                     1    311319 104259158 12844
## - HOUSING_NO_VEHICLE      1    343841 104291680 12844
## - UNINSURED               1    426249 104374088 12845
## - HHCOMP_AGE65            1    440511 104388351 12845
## - AllFlu                  1    543501 104491340 12846
## - MINORITY_NONENGLISH     1    599731 104547571 12847
## - HOUSING_CROWDED         1    611236 104559075 12847
## - HHCOMP_SING_PARENT      1    640092 104587931 12847
## - HHCOMP_DISABILITY       1    718476 104666315 12848
## - MINORITY_MINORITY       1    818484 104766323 12849
## - HHCOMP_AGE17            1    902661 104850500 12850
## - HOUSING_MOBILE          1    954345 104902184 12851
## - HOUSING_MULTIUNIT       1   2147264 106095104 12863
## - pctBiden                1   7827690 111775529 12922
## - SES_INCOME              1  18086185 122034024 13020
##
## Step:  AIC=12846.67
## PERCENT_VAX_0_17 ~ SES_INCOME + HHCOMP_AGE65 + HHCOMP_AGE17 +
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + MMR + HPV + AllFlu + HSA_beds + pctBiden + providers_county
##
##                           Df Sum of Sq       RSS   AIC
## - HSA_beds                 1     68783 104034738 12839
## - MMR                      1     90231 104056187 12840
## <none>                                103965955 12841
## - HOUSING_GROUP_QUARTERS   1    190159 104156115 12841
## - providers_county         1    286047 104252002 12842
## - HPV                      1    325489 104291445 12842
## - HOUSING_NO_VEHICLE       1    326356 104292311 12842
## - HHCOMP_AGE65             1    460396 104426352 12844
## - UNINSURED                1    486979 104452934 12844
## - AllFlu                   1    543360 104509315 12844
## - MINORITY_NONENGLISH      1    624811 104590766 12845
## - HOUSING_CROWDED          1    645549 104611504 12846
## - HHCOMP_SING_PARENT       1    657382 104623337 12846
## - HHCOMP_DISABILITY        1    758753 104724708 12847
## - MINORITY_MINORITY        1    811073 104777028 12847
## - HHCOMP_AGE17             1    910980 104876936 12848
## - HOUSING_MOBILE           1    940009 104905964 12849
## - HOUSING_MULTIUNIT        1   2188349 106154304 12862
## - pctBiden                 1   7840422 111806377 12920
## - SES_INCOME               1  18482140 122448096 13022
##
## Step:  AIC=12845.41
## PERCENT_VAX_0_17 ~ SES_INCOME + HHCOMP_AGE65 + HHCOMP_AGE17 +
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
```

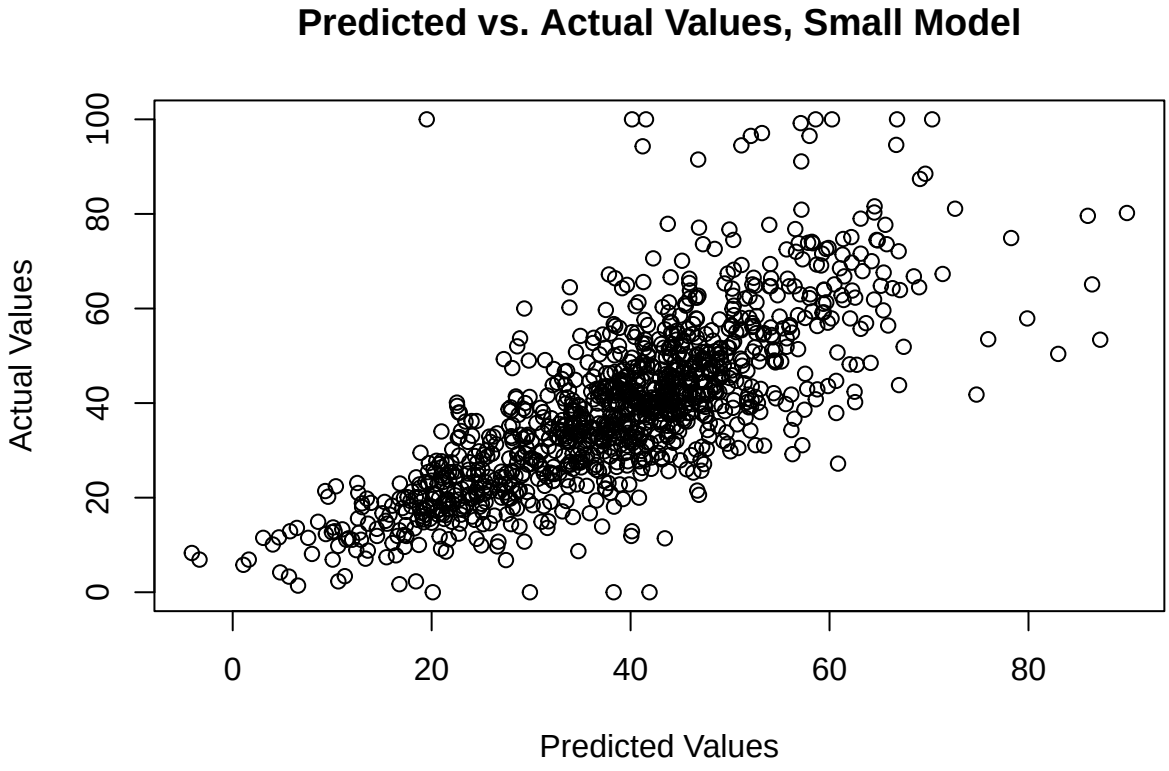
```
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + MMR + HPV + AllFlu + pctBiden + providers_county
##
##                           Df Sum of Sq       RSS   AIC
## - MMR                      1    102614 104137353 12838
## - HOUSING_GROUP_QUARTERS   1    178167 104212905 12839
## <none>                                104034738 12839
## - providers_county         1    229869 104264607 12840
## - HPV                      1    260476 104295214 12840
## - HOUSING_NO_VEHICLE       1    328899 104363637 12841
## - HHCOMP_AGE65             1    449729 104484467 12842
## - UNINSURED                1    490911 104525650 12843
## - MINORITY_NONENGLISH      1    591250 104625988 12844
## - HOUSING_CROWDED          1    625465 104660203 12844
## - HHCOMP_SING_PARENT       1    674302 104709040 12845
## - HHCOMP_DISABILITY        1    726968 104761707 12845
## - MINORITY_MINORITY        1    786652 104821391 12846
## - AllFlu                   1    860150 104894888 12847
## - HHCOMP_AGE17             1    895376 104930114 12847
## - HOUSING_MOBILE           1   1027398 105062137 12848
## - HOUSING_MULTIUNIT        1   2257008 106291746 12861
## - pctBiden                 1   9042856 113077594 12931
## - SES_INCOME               1  18552170 122586908 13021
##
## Step:  AIC=12844.51
## PERCENT_VAX_0_17 ~ SES_INCOME + HHCOMP_AGE65 + HHCOMP_AGE17 +
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + HOUSING_GROUP_QUARTERS +
##     UNINSURED + HPV + AllFlu + pctBiden + providers_county
##
##                           Df Sum of Sq       RSS   AIC
## - HOUSING_GROUP_QUARTERS   1    180524 104317877 12838
## <none>                                104137353 12838
## - HPV                      1    287404 104424756 12840
## - HOUSING_NO_VEHICLE       1    327977 104465330 12840
## - HHCOMP_AGE65             1    497058 104634411 12842
## - UNINSURED                1    527726 104665079 12842
## - MINORITY_NONENGLISH      1    589056 104726409 12843
## - HOUSING_CROWDED          1    623687 104761039 12843
## - HHCOMP_SING_PARENT       1    692247 104829599 12844
## - HHCOMP_DISABILITY        1    696762 104834114 12844
## - MINORITY_MINORITY        1    814097 104951450 12845
## - HHCOMP_AGE17             1    879469 105016822 12846
## - HOUSING_MOBILE           1    970002 105107355 12847
## - providers_county         1   1095293 105232646 12848
## - AllFlu                   1   1106609 105243961 12848
## - HOUSING_MULTIUNIT        1   2202508 106339861 12860
## - pctBiden                 1  17898527 122035880 13014
## - SES_INCOME               1  18709572 122846925 13021
##
## Step:  AIC=12844.45
## PERCENT_VAX_0_17 ~ SES_INCOME + HHCOMP_AGE65 + HHCOMP_AGE17 +
```

```
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + UNINSURED + HPV +
##     AllFlu + pctBiden + providers_county
##
##                       Df Sum of Sq       RSS   AIC
## <none>                             104317877 12838
## - HPV                  1    279387 104597264 12839
## - HOUSING_NO_VEHICLE   1    394631 104712507 12841
## - MINORITY_NONENGLISH  1    581572 104899449 12843
## - HOUSING_CROWDED      1    586210 104904087 12843
## - HHCOMP_AGE65         1    598497 104916373 12843
## - UNINSURED            1    643589 104961466 12843
## - HHCOMP_SING_PARENT   1    699742 105017619 12844
## - HHCOMP_DISABILITY    1    755879 105073756 12844
## - MINORITY_MINORITY    1    847900 105165777 12846
## - providers_county     1    972392 105290269 12847
## - HOUSING_MOBILE       1   1014813 105332690 12847
## - AllFlu               1   1104516 105422393 12848
## - HHCOMP_AGE17         1   1125211 105443088 12848
## - HOUSING_MULTIUNIT    1   2250947 106568824 12860
## - pctBiden             1  17782400 122100277 13013
## - SES_INCOME           1  18529048 122846925 13019
```

```
##
## Call:
## lm(formula = PERCENT_VAX_0_17 ~ SES_INCOME + HHCOMP_AGE65 + HHCOMP_AGE17 +
##     HHCOMP_DISABILITY + HHCOMP_SING_PARENT + MINORITY_MINORITY +
##     MINORITY_NONENGLISH + HOUSING_MULTIUNIT + HOUSING_MOBILE +
##     HOUSING_CROWDED + HOUSING_NO_VEHICLE + UNINSURED + HPV +
##     AllFlu + pctBiden + providers_county, data = model_data,
##     weights = Under_18)
##
## Weighted Residuals:
##     Min      1Q  Median      3Q     Max
## -903.66 -174.04  -18.56  163.03 2615.22
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          38.8465     0.3228 120.350  < 2e-16 ***
## SES_INCOME            7.2733     0.5206  13.972  < 2e-16 ***
## HHCOMP_AGE65         -1.3713     0.5461  -2.511 0.012181 *
## HHCOMP_AGE17         -1.8777     0.5454  -3.443 0.000597 ***
## HHCOMP_DISABILITY    -1.6519     0.5854  -2.822 0.004860 **
## HHCOMP_SING_PARENT   -1.1487     0.4231  -2.715 0.006729 **
## MINORITY_MINORITY     1.9939     0.6671   2.989 0.002863 **
## MINORITY_NONENGLISH   1.4490     0.5854   2.475 0.013464 *
## HOUSING_MULTIUNIT    -2.3980     0.4924  -4.870 1.28e-06 ***
## HOUSING_MOBILE        1.0925     0.3341   3.270 0.001110 **
## HOUSING_CROWDED      -1.1764     0.4734  -2.485 0.013099 *
## HOUSING_NO_VEHICLE    0.9933     0.4872   2.039 0.041690 *
## UNINSURED            -1.2937     0.4968  -2.604 0.009341 **
## HPV                   1.0049     0.5857   1.716 0.086512 .
## AllFlu                1.6234     0.4759   3.411 0.000670 ***
```

```
## pctBiden              6.5333     0.4773  13.687  < 2e-16 ***
## providers_county     -1.2093     0.3778  -3.201 0.001410 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 308.1 on 1099 degrees of freedom
## Multiple R-squared:  0.6357, Adjusted R-squared:  0.6304
## F-statistic: 119.8 on 16 and 1099 DF,  p-value: < 2.2e-16
```

```
##                Variable    Weight
## 1           SES_INCOME 7.2732988
## 2             pctBiden 6.5332702
## 3    MINORITY_MINORITY 1.9938809
## 4                AllFlu 1.6233967
## 5  MINORITY_NONENGLISH 1.4489727
## 6       HOUSING_MOBILE 1.0924814
## 7                  HPV 1.0048644
## 8   HOUSING_NO_VEHICLE 0.9933149
## 9   HHCOMP_SING_PARENT 1.1487451
## 10     HOUSING_CROWDED 1.1763547
## 11    providers_county 1.2093257
## 12           UNINSURED 1.2936773
## 13        HHCOMP_AGE65 1.3712781
## 14   HHCOMP_DISABILITY 1.6519282
## 15        HHCOMP_AGE17 1.8777205
## 16   HOUSING_MULTIUNIT 2.3980421
```



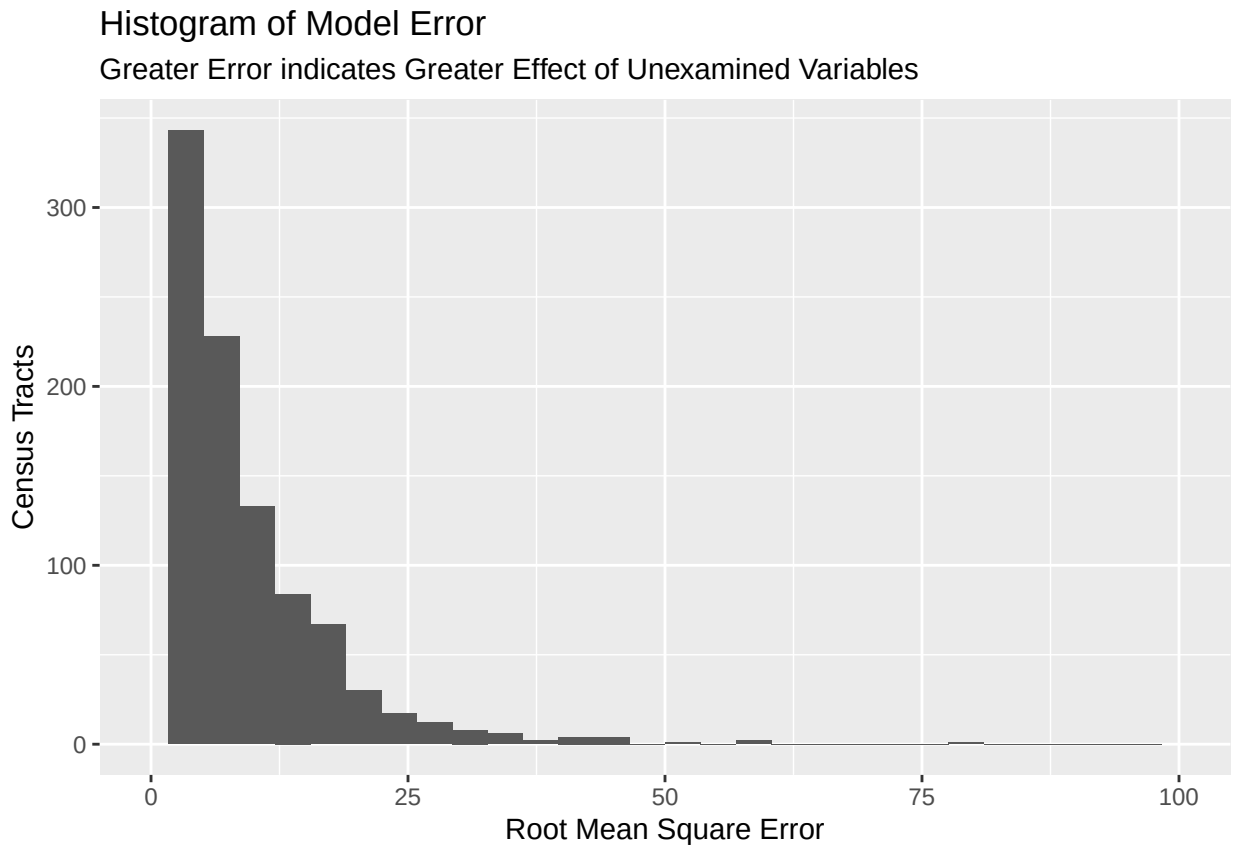
Then we make a map because Shannon asked me to

```
model_fit_graph_df <- model_data %>%
  mutate(model_prediction = predict(model_small)) %>%
  rowwise() %>%
  mutate(model_error = sqrt(
    mean(
      (PERCENT_VAX_0_17 - model_prediction)^2
      )
         )
         ) %>%
  ungroup() %>%
  left_join(Colorado_data %>% select(TractID,geometry))
```

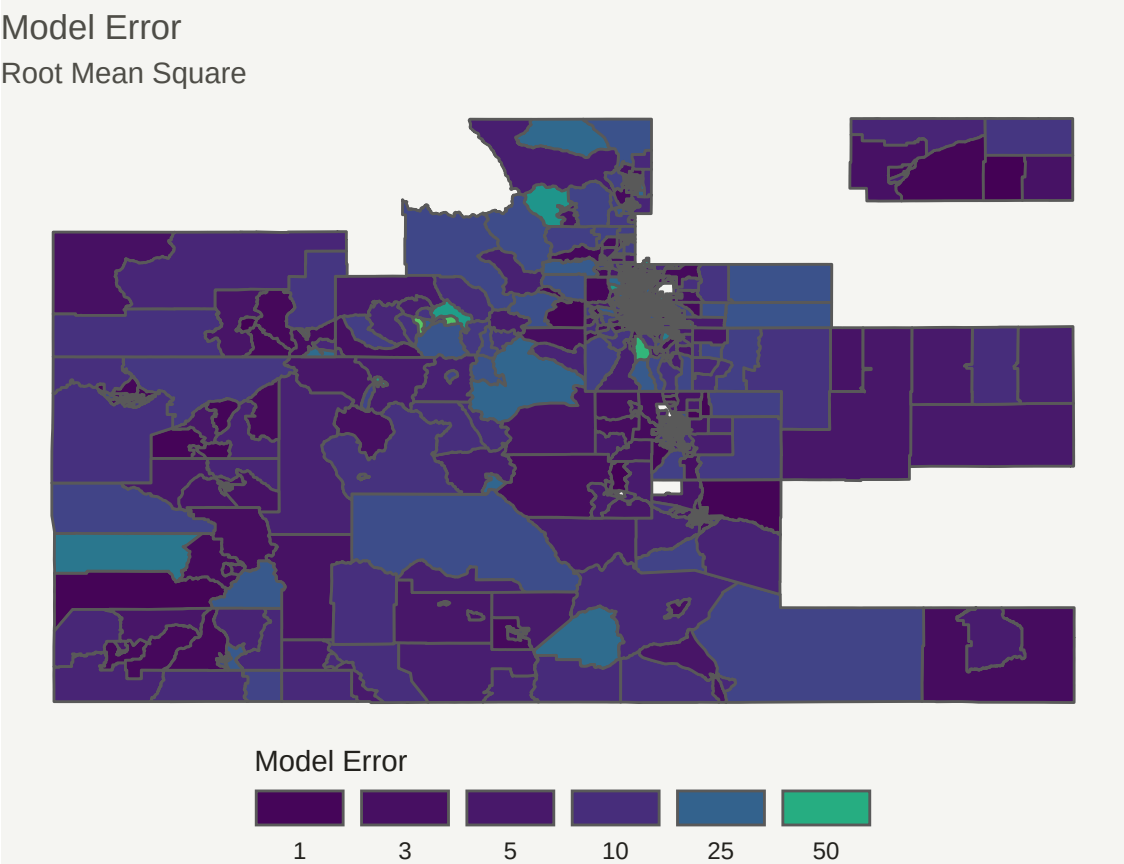
```
## Joining, by = c("TractID", "geometry")
```

```
model_fit_graph_df  %>%
  ggplot(aes(x=`model_error`)
         ) +
  geom_histogram(bins = 30) +
  ylab("Census Tracts")+
  xlab("Root Mean Square Error")+
  labs(title = "Histogram of Model Error",
       subtitle = "Greater Error indicates Greater Effect of Unexamined Variables") +
  xlim(0,100)
```

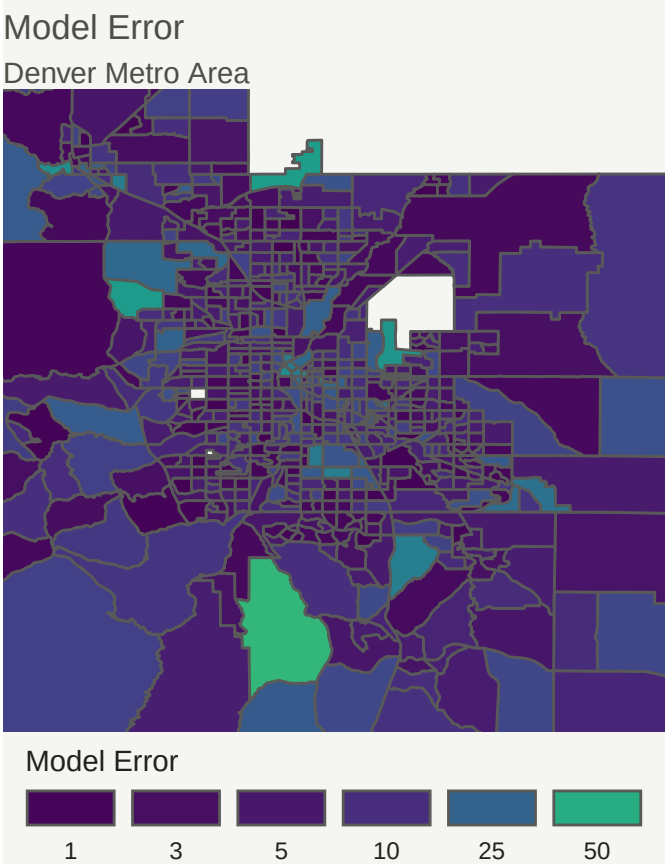
```
## Warning: Removed 2 rows containing missing values (geom_bar).
```




```
# https://r-graph-gallery.com/327-chloropleth-map-from-geojson-with-ggplot2.html
model_fit_graph_df %>%
  ggplot(aes(fill = `model_error`,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(breaks=c(1,3,5,10,25,50), name="Model Error", guide = guide_legend( keyheight = u
  labs(title = "Model Error",
       subtitle = "Root Mean Square") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  )
```




```
model_fit_graph_df %>%
  ggplot(aes(fill = `model_error`,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(breaks=c(1,3,5,10,25,50), name="Model Error", guide = guide_legend( keyheight = u
  labs(title = "Model Error",
       subtitle = "Root Mean Square") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  ) +
  xlim(-105.332,-104.559)+
  ylim(39.321,40.072)+
  labs(title = "Model Error",
       subtitle = "Denver Metro Area")
```

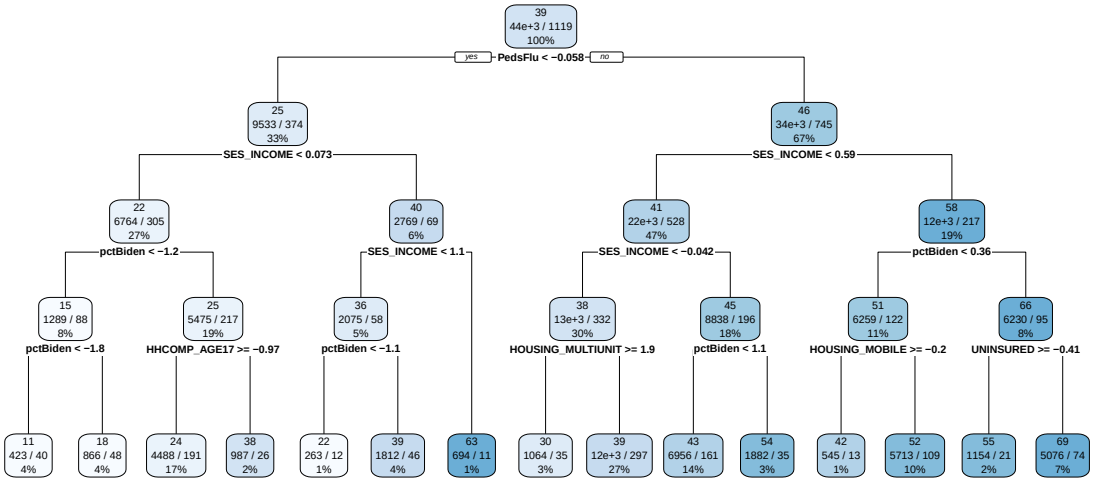


Random Forest

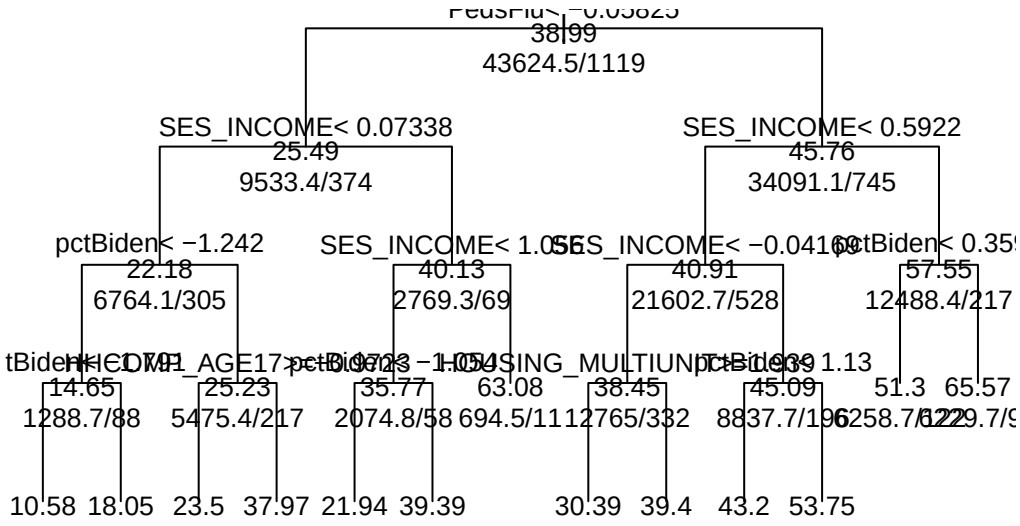
This does machine learning! Suffice to say that it is a neat mthodology.

Importance graph: * Mean Decrease Accuracy (%IncMSE) - This shows how much our model accuracy decreases if we leave out that variable.
* Mean Decrease Gini (IncNodePurity) - This is a measure of variable importance based on the Gini impurity index used for the calculating the splits in trees.

The higher the value of mean decrease accuracy or mean decrease gini score, the higher the importance of the variable to our model.


39
44e+3 / 1119
100%
yes
PedsFlu < −0.058
no
25
9533 / 374
33%
46
34e+3 / 745
67%
SES_INCOME < 0.073
SES_INCOME < 0.59
22
6764 / 305
27%
40
2769 / 69
6%
41
22e+3 / 528
47%
58
12e+3 / 217
19%
pctBiden < −1.2
SES_INCOME < 1.1
SES_INCOME < −0.042
pctBiden < 0.36
15
1289 / 88
8%
25
5475 / 217
19%
36
2075 / 58
5%
38
13e+3 / 332
30%
45
8838 / 196
18%
51
6259 / 122
11%
66
6230 / 95
8%
pctBiden < −1.8
HHCOMP_AGE17 >= −0.97
pctBiden < −1.1
HOUSING_MULTIUNIT >= 1.9
pctBiden < 1.1
HOUSING_MOBILE >= −0.2
UNINSURED >= −0.41
11
423 / 40
4%
18
866 / 48
4%
24
4488 / 191
17%
38
987 / 26
2%
22
263 / 12
1%
39
1812 / 46
4%
63
694 / 11
1%
30
1064 / 35
3%
39
12e+3 / 297
27%
43
6956 / 161
14%
54
1882 / 35
3%
42
545 / 13
1%
52
5713 / 109
10%
55
1154 / 21
2%
69
5076 / 74
7%

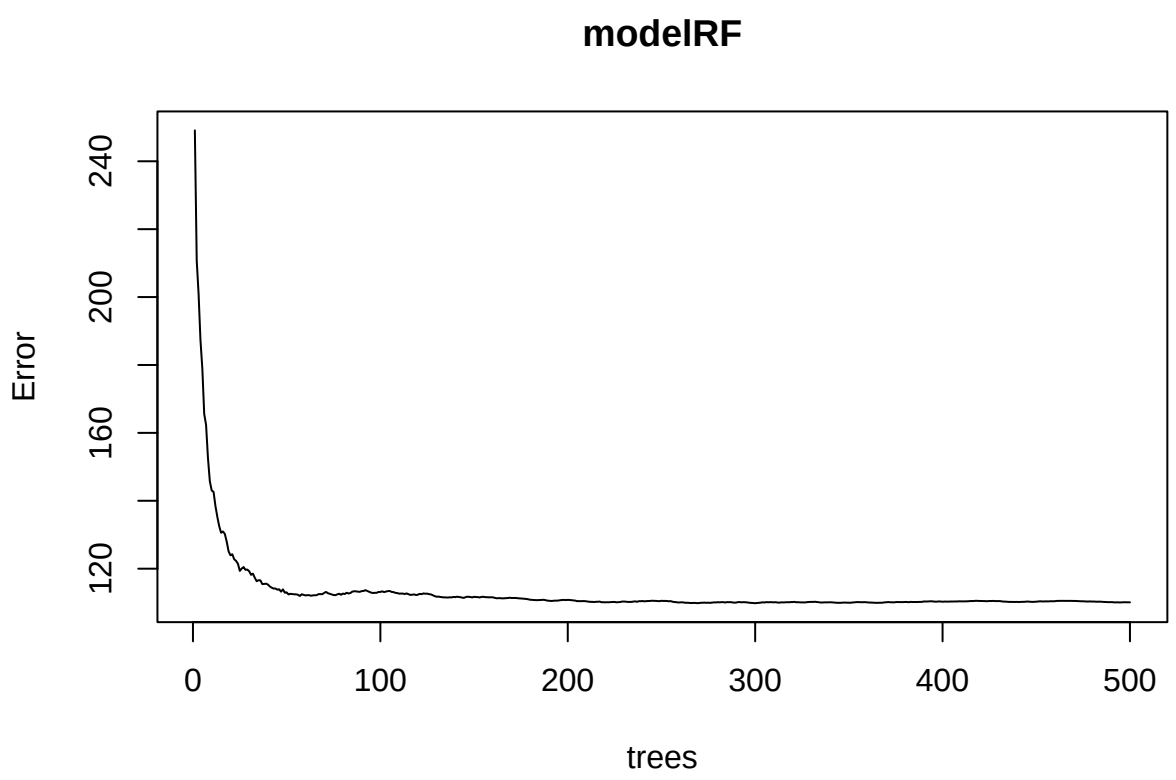
Pruned Classification Tree For Sales



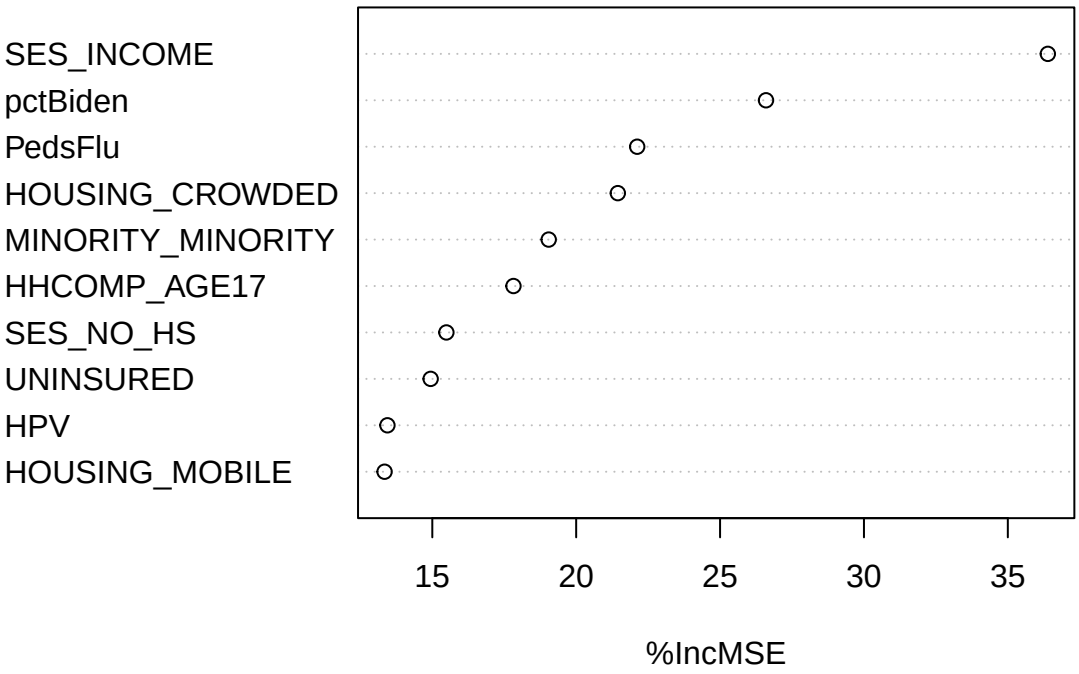

```
##
## Call:
##  randomForest(formula = PERCENT_VAX_0_17 ~ SES_BELOW_POV + SES_UNEMPLOYED +      SES_INCOME + SES_NO
##                Type of random forest: regression
##                      Number of trees: 500
## No. of variables tried at each split: 7
##
##           Mean of squared residuals: 110.0843
##                     % Var explained: 63.71

##                 Length Class  Mode
## call               5   -none- call
## type               1   -none- character
## predicted       1119   -none- numeric
## mse              500   -none- numeric
## rsq              500   -none- numeric
## oob.times       1119   -none- numeric
## importance        46   -none- numeric
## importanceSD      23   -none- numeric
## localImportance    0   -none- NULL
## proximity          0   -none- NULL
## ntree              1   -none- numeric
## mtry               1   -none- numeric
## forest            11   -none- list
## coefs              0   -none- NULL
```

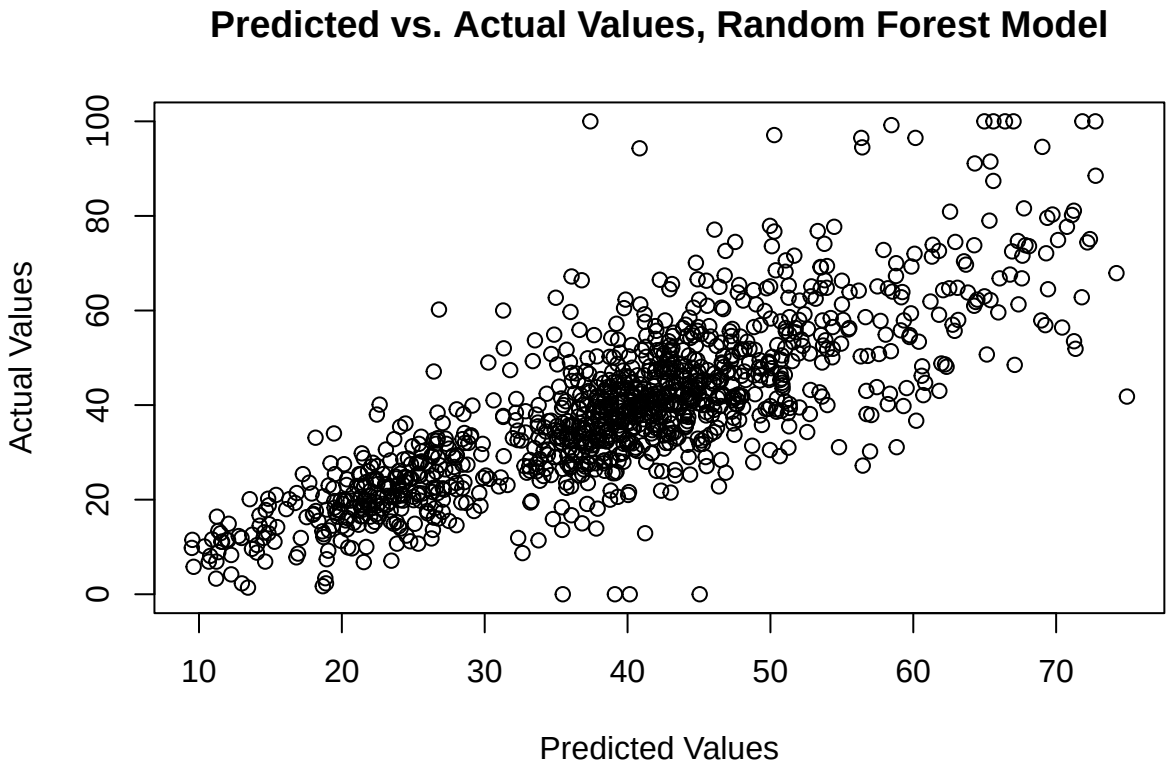
```
## y               1119   -none- numeric
## test               0   -none- NULL
## inbag              0   -none- NULL
## terms              3   terms  call
```

```
##                          %IncMSE IncNodePurity
## SES_BELOW_POV          10.017393     6997.773
## SES_UNEMPLOYED          4.237680     5832.242
## SES_INCOME             36.389909    59507.059
## SES_NO_HS              15.489738    12122.064
## HHCOMP_AGE65            6.952244     8807.154
## HHCOMP_AGE17           17.820638    16108.790
## HHCOMP_DISABILITY      12.739219    12273.806
## HHCOMP_SING_PARENT     11.472807    11849.725
## MINORITY_MINORITY      19.045243    10213.079
## MINORITY_NONENGLISH    12.852907     5672.378
## HOUSING_MULTIUNIT      10.453171    10951.988
## HOUSING_MOBILE         13.341532     6026.121
## HOUSING_CROWDED        21.446972    11435.064
## HOUSING_NO_VEHICLE      7.996709     6713.954
## HOUSING_GROUP_QUARTERS  3.590869     4577.707
## UNINSURED              14.942618    10813.157
## MMR                    13.241046     5359.992
## HPV                    13.441012    20604.604
## PedsFlu                22.120073    38888.450
## AllFlu                 10.806066     8049.208
## HSA_beds                8.932963     8854.865
## pctBiden               26.596656    41437.291
## providers_county       11.641665     8359.961
```



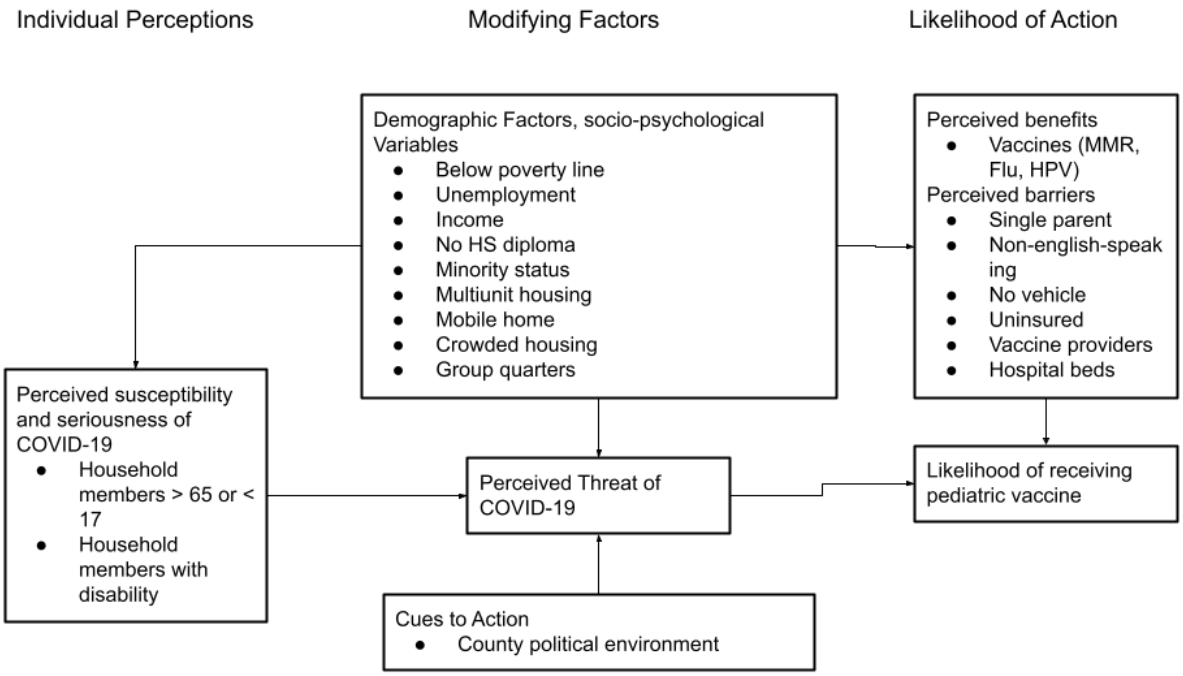

```
##                           %IncMSE IncNodePurity
## SES_BELOW_POV           10.017393      6997.773
## SES_UNEMPLOYED           4.237680      5832.242
## SES_INCOME              36.389909     59507.059
## SES_NO_HS               15.489738     12122.064
## HHCOMP_AGE65             6.952244      8807.154
## HHCOMP_AGE17            17.820638     16108.790
## HHCOMP_DISABILITY       12.739219     12273.806
## HHCOMP_SING_PARENT      11.472807     11849.725
## MINORITY_MINORITY       19.045243     10213.079
## MINORITY_NONENGLISH     12.852907      5672.378
## HOUSING_MULTIUNIT       10.453171     10951.988
## HOUSING_MOBILE          13.341532      6026.121
## HOUSING_CROWDED         21.446972     11435.064
## HOUSING_NO_VEHICLE       7.996709      6713.954
## HOUSING_GROUP_QUARTERS   3.590869      4577.707
## UNINSURED               14.942618     10813.157
## MMR                     13.241046      5359.992
## HPV                     13.441012     20604.604
## PedsFlu                 22.120073     38888.450
## AllFlu                  10.806066      8049.208
## HSA_beds                 8.932963      8854.865
## pctBiden                26.596656     41437.291
## providers_county        11.641665      8359.961
```


Variable Importance, Random Forest Model
SES_INCOME
pctBiden
PedsFlu
HOUSING_CROWDED
MINORITY_MINORITY
HHCOMP_AGE17
SES_NO_HS
UNINSURED
HPV
HOUSING_MOBILE
15
20
25
30
35
%IncMSE


Predicted vs. Actual Values, Random Forest Model
Actual Values
0
20
40
60
80
100
10
20
30
40
50
60
70
Predicted Values

Causal model

The next thing we do, after all that machine learning, is to actually implement some kind of model. Here's the model:



From: The Health Belief Model: A Decade Later, Janz, Becker, Health Education Quarterly 1984

And that's it!

Well, it for now. But I think it's really something to go on. Other things we can do would be to:

- List the highest and lowest performing census tracts, counties, etc.
- For each tract, figure out what is the most impactful factor
- And maybe some other stuff.

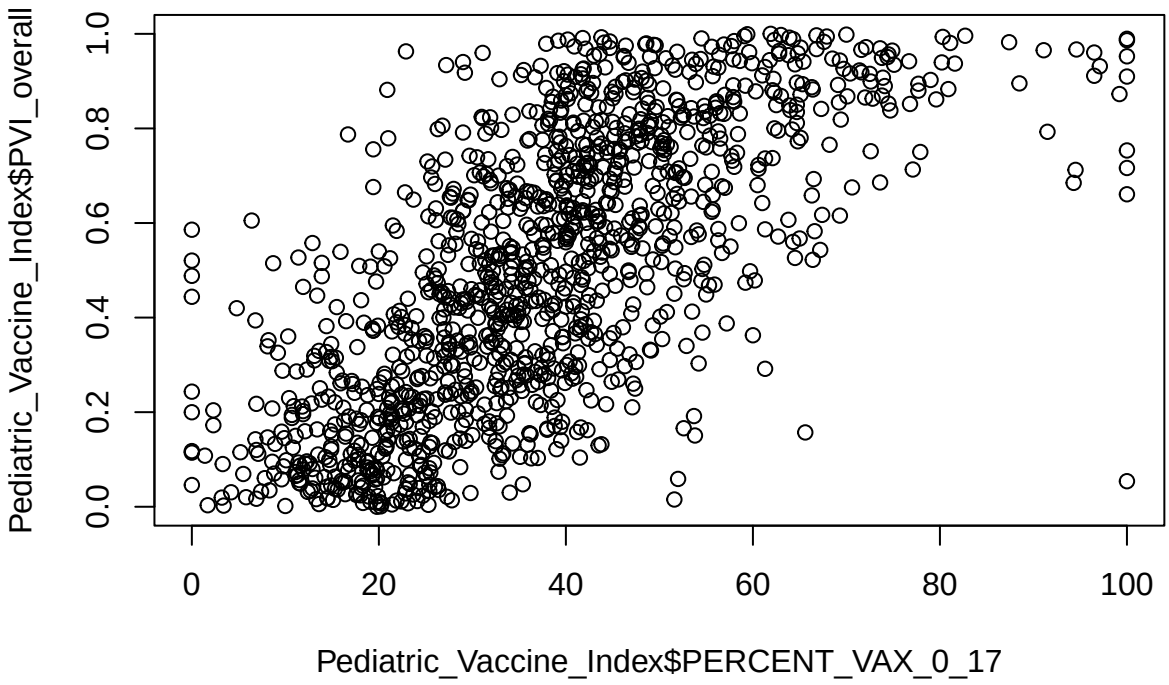
```
Variable_pct <- Colorado_data_mr %>%
  select(TractID,geometry,PERCENT_VAX_0_17,

         HPV,
AllFlu,
pctBiden,

HHCOMP_AGE65,
HHCOMP_AGE17,
HHCOMP_DISABILITY,
```

```
HHCOMP_SING_PARENT,
HOUSING_NO_VEHICLE,
UNINSURED,
providers_county,

SES_INCOME,
MINORITY_MINORITY,
MINORITY_NONENGLISH,
HOUSING_MULTIUNIT,
HOUSING_MOBILE,
HOUSING_CROWDED

) %>%
  mutate(across(HPV:HOUSING_CROWDED,
                ~ percent_rank(.x),
                .names = "{.col}_pct"))  %>%
  ungroup()

#Beliefs about vaccine
Belief_pct <- Variable_pct %>%

mutate(Belief_pct = percent_rank(
  HPV_pct+AllFlu_pct+pctBiden_pct
)) %>%
  select(TractID,Belief_pct) %>%
  ungroup()

# Perceeived vulnerability
Perc_Vuln_pct <- Variable_pct %>%

mutate(Perc_Vuln_pct  = percent_rank(
  -1.3713*(HHCOMP_AGE65)+
-1.8777*(HHCOMP_AGE17)+
-1.6519*(HHCOMP_DISABILITY)
)) %>%
  select(TractID,Perc_Vuln_pct)

#Healthcare access
HC_access_pct <- Variable_pct %>%


mutate(HC_access_pct = percent_rank(
  -1.1487*(HHCOMP_SING_PARENT)+
0.9933*(HOUSING_NO_VEHICLE)+
-1.2937*(UNINSURED)+
-1.2093*(providers_county)
)
) %>%
select(HC_access_pct,TractID)

#Socioeconomic disadvantage
SE_Disadv_pct <- Variable_pct %>%
```

```
mutate(SE_Disadv_pct = percent_rank(
  7.2733*SES_INCOME+
1.9939*MINORITY_MINORITY+
.4490*MINORITY_NONENGLISH+
-2.3980*HOUSING_MULTIUNIT+
1.0925*HOUSING_MOBILE+
-1.1764*HOUSING_CROWDED
)) %>%
select(SE_Disadv_pct,TractID)

theme_vars <- Variable_pct %>%
  select(geometry,TractID,PERCENT_VAX_0_17,ends_with("_pct")) %>%
  left_join(Belief_pct, by = "TractID") %>%
  left_join(Perc_Vuln_pct, by = "TractID") %>%
  left_join(HC_access_pct, by = "TractID") %>%
  left_join(SE_Disadv_pct, by = "TractID")

themes_weight_model <- lm(PERCENT_VAX_0_17 ~
                            Belief_pct +
                            Perc_Vuln_pct+
                            HC_access_pct+
                            SE_Disadv_pct,
                          data = theme_vars)

print(
  summary(
    themes_weight_model
  )
)
```

```
##
## Call:
## lm(formula = PERCENT_VAX_0_17 ~ Belief_pct + Perc_Vuln_pct +
##     HC_access_pct + SE_Disadv_pct, data = theme_vars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -41.426  -8.082  -1.062   7.064  81.093
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.0582     0.9564  11.563  < 2e-16 ***
## Belief_pct     21.5773     1.4705  14.673  < 2e-16 ***
## Perc_Vuln_pct   7.8378     1.4234   5.507 4.45e-08 ***
## HC_access_pct   1.7597     1.5069   1.168    0.243
## SE_Disadv_pct  23.6948     1.6411  14.438  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.86 on 1237 degrees of freedom
## Multiple R-squared:  0.4735, Adjusted R-squared:  0.4718
## F-statistic: 278.1 on 4 and 1237 DF,  p-value: < 2.2e-16
```

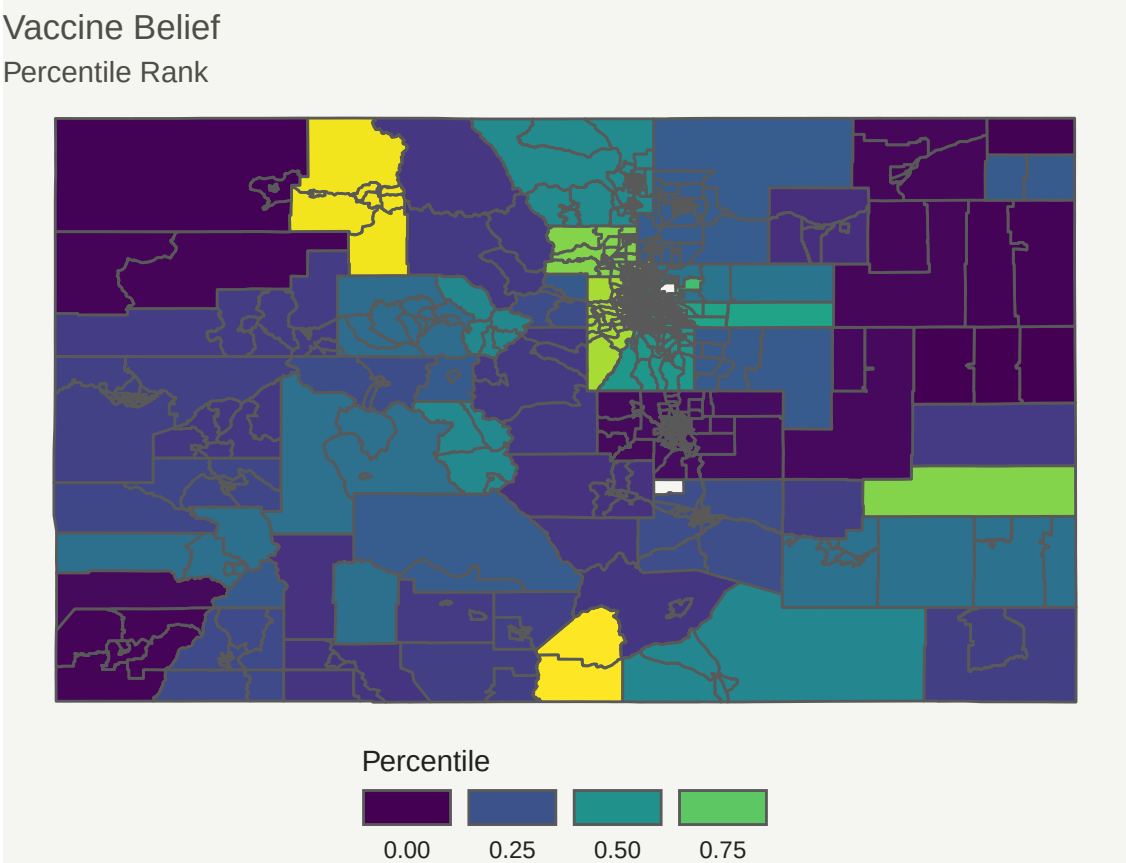
```
Pediatric_Vaccine_Index <- Variable_pct %>%
  select(geometry,TractID,PERCENT_VAX_0_17,ends_with("_pct")) %>%
  left_join(Belief_pct, by = "TractID") %>%
  left_join(Perc_Vuln_pct, by = "TractID") %>%
  left_join(HC_access_pct, by = "TractID") %>%
  left_join(SE_Disadv_pct, by = "TractID") %>%
  mutate(PVI_overall = percent_rank(
  21.5773*Belief_pct+
  7.8378*Perc_Vuln_pct+
  1.7597*HC_access_pct+
  23.6948*SE_Disadv_pct
  ))


plot(Pediatric_Vaccine_Index$PERCENT_VAX_0_17,Pediatric_Vaccine_Index$PVI_overall)
```




```
write_csv(Pediatric_Vaccine_Index %>% select(-geometry),"data/Pediatric_Vaccine_Index.csv")
```

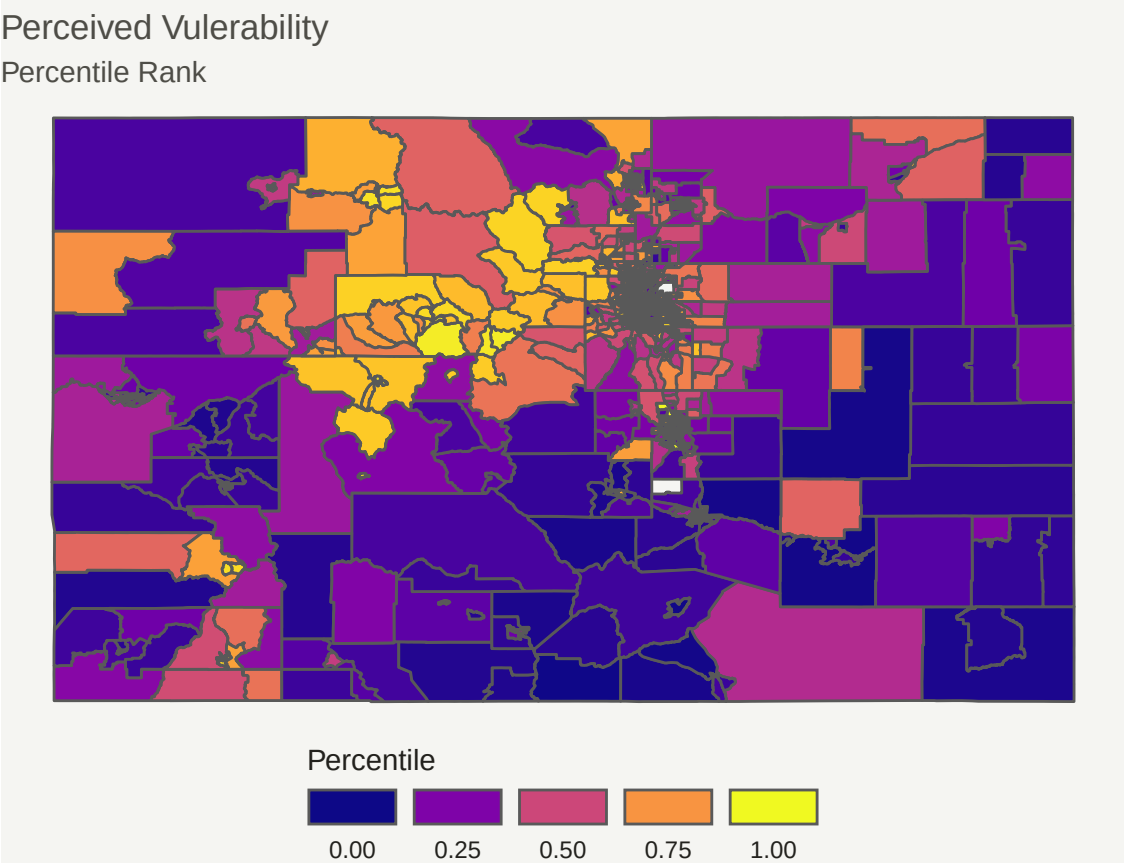
```
Pediatric_Vaccine_Index %>%
  ggplot(aes(fill = Belief_pct,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(option = "D",
      breaks=c(0,.25,.5,.75,1), name="Percentile", guide = guide_legend( keyheight = unit(5, units = "m
  labs(title = "Vaccine Belief",
```

```
      subtitle = "Percentile Rank") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  )
```



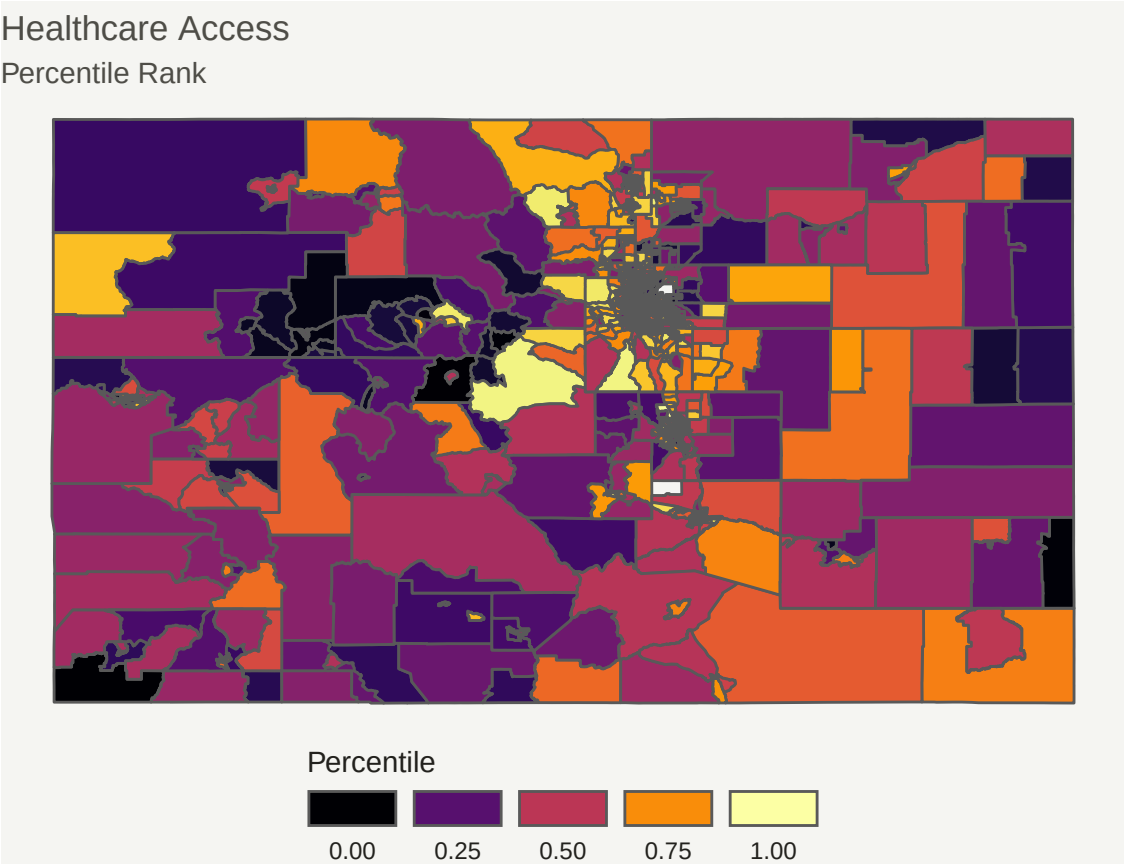

```
Pediatric_Vaccine_Index %>%
  ggplot(aes(fill = Perc_Vuln_pct,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(option = "C",
      breaks=c(0,.25,.5,.75,1), name="Percentile", guide = guide_legend( keyheight = unit(5, units = "mm
  labs(title = "Perceived Vulerability",
       subtitle = "Percentile Rank") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
```

```
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  )
```



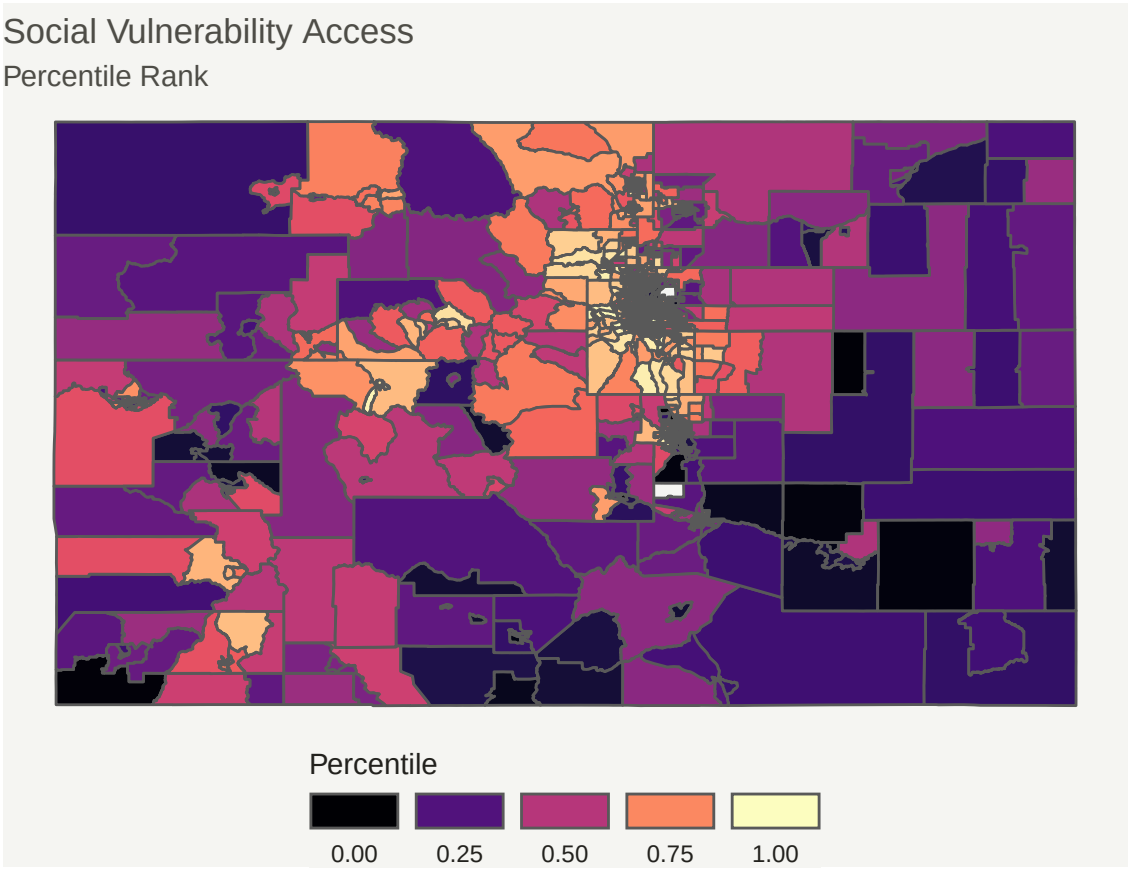

```
Pediatric_Vaccine_Index %>%
  ggplot(aes(fill = HC_access_pct,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(option = "B",
      breaks=c(0,.25,.5,.75,1), name="Percentile", guide = guide_legend( keyheight = unit(5, units = "m
  labs(title = "Healthcare Access",
       subtitle = "Percentile Rank") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
```

```
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  )
```

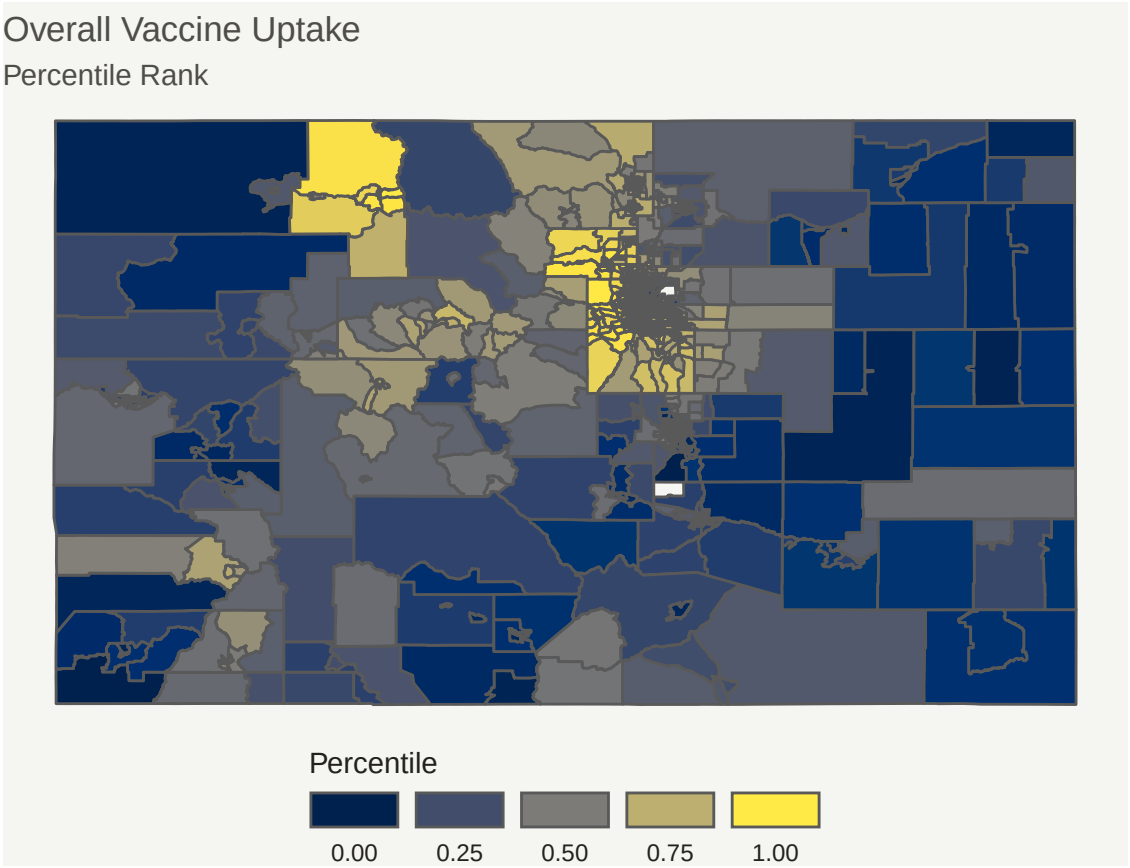



```
Pediatric_Vaccine_Index %>%
  ggplot(aes(fill = SE_Disadv_pct,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(option = "A",
      breaks=c(0,.25,.5,.75,1), name="Percentile", guide = guide_legend( keyheight = unit(5, units = "m
  labs(title = "Social Vulnerability Access",
       subtitle = "Percentile Rank") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
```

```
  )
```




```
Pediatric_Vaccine_Index %>%
  ggplot(aes(fill = PVI_overall,geometry = geometry)) +
  geom_sf()+
    theme_void() +
    scale_fill_viridis(option = "E",
      breaks=c(0,.25,.5,.75,1), name="Percentile", guide = guide_legend( keyheight = unit(5, units = "m
  labs(title = "Overall Vaccine Uptake",
       subtitle = "Percentile Rank") +
  theme(
    text = element_text(color = "#22211d"),
    plot.background = element_rect(fill = "#f5f5f2", color = NA),
    panel.background = element_rect(fill = "#f5f5f2", color = NA),
    legend.background = element_rect(fill = "#f5f5f2", color = NA),

    plot.title = element_text(color = "#4e4d47"),
    plot.subtitle = element_text(color = "#4e4d47"),
    plot.caption = element_text( size=12, color = "#4e4d47"),

    legend.position = "bottom"
  )
```




```
# Colorado_data_mr %>%
#   select(PERCENT_VAX_0_17,
#
#          HPV,
# AllFlu,
# pctBiden,
#
# HHCOMP_AGE65,
# HHCOMP_AGE17,
# HHCOMP_DISABILITY,
#
# HHCOMP_SING_PARENT,
# HOUSING_NO_VEHICLE,
# UNINSURED,
# providers_county,
#
# SES_INCOME,
# MINORITY_MINORITY,
# MINORITY_NONENGLISH,
# HOUSING_MULTIUNIT,
# HOUSING_MOBILE,
# HOUSING_CROWDED
#
# ) %>% plot()
```